

CONSEQUENTIALISM AND INTEGRITY*

BERNARD WILLIAMS

1. THE STRUCTURE OF CONSEQUENTIALISM

NO one can hold that everything, of whatever category, that has value, has it in virtue of its consequences. If that were so, one would just go on for ever, and there would be an obviously hopeless regress. That regress would be hopeless even if one takes the view, which is not an absurd view, that although men set themselves ends and work towards them, it is very often not really the supposed end, but the effort towards it on which they set value—that they travel, not really in order to arrive (for as soon as they have arrived they set out for somewhere else), but rather they choose somewhere to arrive, in order to travel. Even on that view, not everything would have consequential value; what would have non-consequential value would in fact be travelling, even though people had to think of travelling as having the consequential value, and something else—the destination—the non-consequential value.

If not everything that has value has it in virtue of consequences, then presumably there are some types of thing which have non-consequential value, and also some particular things that have such value because they are instances of those types. Let us say, using a traditional term, that anything that has that sort of value, has *intrinsic* value.¹ I take it to be the central idea of consequen-

Bernard Williams, from *Utilitarianism: For and Against*, ed. Smart and Williams (Cambridge University Press, 1973), pp. 82–118. Reprinted by permission of the publisher.

* This is not the title of the original printing.

¹ The terminology of things 'being valuable', 'having intrinsic value', etc., is not meant to beg any questions in general value-theory. Non-cognitive theories, such as Smart's, should be able to recognize the distinctions made here.

tialism that the only kind of thing that has intrinsic value is states of affairs, and that anything else that has value has it because it conduces to some intrinsically valuable state of affairs.

How much, however, does this say? Does it succeed in distinguishing consequentialism from anything else? The trouble is that the term 'state of affairs' seems altogether too permissive to exclude anything: may not the obtaining of absolutely anything be represented formally as a state of affairs? A Kantian view of morality, for instance, is usually thought to be opposed to consequentialism, if any is; at the very least, if someone were going to show that Kantianism collapsed into consequentialism, it should be the product of a long and unobvious argument, and not just happen at the drop of a definition. But on the present account it looks as though Kantianism can be made instantly into a kind of consequentialism—a kind which identifies the states of affairs that have intrinsic value (or at least intrinsic moral value) as those that consist of actions being performed for duty's sake.² We need something more to our specification if it is to be the specification of anything distinctly consequentialist.

The point of saying that consequentialism ascribes intrinsic value to states of affairs is rather to *contrast* states of affairs with other candidates for having such value: in particular, perhaps, actions. A distinctive mark of consequentialism might rather be this, that it regards the value of actions as always consequential (or, as we may more generally say, derivative), and not intrinsic. The value of actions would then lie in their causal properties, of producing valuable states of affairs; or if they did not derive their value in this simple way, they would derive it in some more roundabout way, as for instance by being expressive of some motive, or in accordance with some rule, whose operation in society conduces to desirable states of affairs. (The lengths to which such indirect derivations can be taken without wrecking the point of consequentialism is something we shall be considering later.³)

² A point noted by Smart, p. 13. [All of Williams's references to J. J. C. Smart are to Smart's essay 'An Outline of a System of Utilitarian Ethics', in *Utilitarianism: For and Against*, pp. 3–74—Ed.]

³ [Williams is referring here to a section of his essay that is not reprinted in this volume—Ed.]

To insist that what has intrinsic value is states of affairs and not actions seems to come near an important feature of consequentialism. Yet it may be that we have still not hit exactly what we want, and that the restriction is now too severe. Surely *some* actions, compatibly with consequentialism, might have intrinsic value? This is a question which has a special interest for utilitarianism, that is to say, the form of consequentialism concerned particularly with happiness. Traditionally utilitarians have tended to regard happiness or, again, pleasure, as experiences or sensations which were related to actions and activity as effect to cause; and, granted that view, utilitarianism will indeed see the value of all action as derivative, intrinsic value being reserved for the experiences of happiness. But that view of the relations between action and either pleasure or happiness is widely recognized to be inadequate. To say that a man finds certain actions or activity pleasant, or that they make him happy, or that he finds his happiness in them, is certainly not always to say that they induce certain sensations in him, and in the case of happiness, it is doubtful whether that is ever what is meant. Rather it means such things (among others) as that he enjoys doing these things for their own sake. It would trivialize the discussion of utilitarianism to tie it by definition to inadequate conceptions of happiness or pleasure, and we must be able to recognize as versions of utilitarianism those which, as most modern versions do, take as central some notion such as *satisfaction*, and connect that criterially with such matters as the activities which a man will freely choose to engage in. But the activities which a man engages in for their own sake are activities in which he finds intrinsic value. So any specification of consequentialism which logically debar action or activity from having intrinsic value will be too restrictive even to admit the central case, utilitarianism, so soon as that takes on a more sophisticated and adequate conception of its basic value of happiness.

So far then, we seem to have one specification of consequentialism which is too generous to exclude anything, and another one which is too restrictive to admit even the central case. These difficulties arise from either admitting without question actions among desirable states of affairs, or blankly excluding all actions from the state of affairs category. This suggests that we shall do

better by looking at the interrelations between states of affairs and actions.

It will be helpful, in doing this, to introduce the notion of the *right* action for an agent in given circumstances. I take it that in any form of direct consequentialism, and certainly in act-utilitarianism, the notion of the right action in given circumstances is a maximizing notion:⁴ the right action is that which out of the actions available to the agent brings about or represents the highest degree of whatever it is the system in question regards as intrinsically valuable—in the central case, utilitarianism, this is of course happiness. In this argument, I shall confine myself to direct consequentialism, for which 'right action' is unqualifiedly a maximizing notion.

The notion of the right action as that which, of the possible alternatives, maximizes the good (where this embraces, in unfavourable circumstances, minimizing the bad), is an objective notion in this sense, that it is perfectly possible for an agent to be ignorant or mistaken, and non-culpably ignorant or mistaken, about what is the right action in the circumstances. Thus the assessment by others of whether the agent did, in this sense, do the right thing, is not bounded by the agent's state of knowledge at the time, and the claim that he did the wrong thing is compatible with recognizing that he did as well as anyone in his state of knowledge could have done.⁵ It might be suggested that, contrary to this, we have already imported the subjective conditions of action in speaking of the best of the actions *available to him*: if he is ignorant or misinformed, then the actions which might seem to us available to him were not in any real sense available. But this would be an exaggeration; the notion of availability imports some, but not all, kinds of subjective condition. Over and above the question of actions which, granted his situation and powers, were physically not available to him, we might perhaps add that a course of action was not really available to an agent if his historical, cultural, or psychological situation was such that it could not possibly occur to him. But it is scarcely reasonable to extend the notion of

⁴ Cf. Smart's definition, p. 45.

⁵ In Smart's terminology, the 'rational thing': pp. 46-7.

unavailability to actions which merely did not occur to him; and surely absurd to extend it to actions which did occur to him, but where he was misinformed about their consequences.

If then an agent does the right thing, he does the best of the alternatives available to him (where that, again, embraces the least bad: we shall omit this rider from now on). Standardly, the action will be right in virtue of its causal properties, of maximally conducing to good states of affairs. Sometimes, however, the relation of the action to the good state of affairs may not be that of cause to effect—the good state of affairs may be constituted, or partly constituted, by the agent's doing that act (as when under utilitarianism he just enjoys doing it, and there is no project available to him more productive of happiness for him or anyone else).

Although this may be so under consequentialism, there seems to be an important difference between this situation and a situation of an action's being right for some non-consequentialist reason, as for instance under a Kantian morality. This difference might be brought out intuitively by saying that for the consequentialist, even a situation of this kind in which the action itself possesses intrinsic value is one in which the rightness of the act is derived from the goodness of a certain state of affairs—the act is right *because* the state of affairs which consists in its being done is better than any other state of affairs accessible to the agent; whereas for the non-consequentialist it is sometimes, at least, the other way round, and a state of affairs which is better than the alternatives is so because it consists of the right act being done. This intuitive description of the difference has something in it, but it needs to be made more precise.

We can take a step towards making it more precise, perhaps, in the following way. Suppose *S* is some particular concrete situation. Consider the statement, made about some particular agent

In *S*, he did the right thing in doing *A*. (1)

For consequentialists, (1) implies a statement of the form

The state of affairs *P* is better than any other state of affairs accessible to him. (2)

Here a state of affairs being 'accessible' to an agent means that it is

a state of affairs which is the consequence of, or is constituted by, his doing an act available to him (for that, see above); and *P* is a state of affairs accessible to him only in virtue of his doing *A*.⁶

Now in the exceptional case where it is just his doing *A* which carries the intrinsic value, we get for (2)

The state of affairs which consists in his doing *A* is better than any other state of affairs accessible to him. (3)

It was just the possibility of this sort of case which raised the difficulty of not being able to distinguish between a sophisticated consequentialism and non-consequentialism. The question thus is: if (3) is what we get for consequentialism in this sort of case, is it what a non-consequentialist would regard as implied by (1)? If so, we still cannot tell the difference between them. But the answer in fact seems to be 'no'.

There are two reasons for this. One reason is that a non-consequentialist, though he must inevitably be able to attach a sense to (1), does not have to be able to attach a sense to (3) at all, while the consequentialist, of course, attaches a sense to (1) only because he attaches a sense to (3). Although the non-consequentialist is concerned with right actions—such as the carrying out of promises—he may have no general way of comparing states of affairs from a moral point of view at all. Indeed, we shall see later, and in greater depth than these schematic arguments allow, that the emphasis on the necessary comparability of situations is a peculiar feature of consequentialism in general, and of utilitarianism in particular.

A different kind of reason emerges if we suppose that the non-consequentialist does admit, in general, comparison between states of affairs. Thus, we might suppose that some non-consequentialist would consider it a better state of things in which more, rather than fewer, people kept their promises, and kept them for non-consequentialist reasons. Yet consistently with that he could accept, in a particular case, all of the following: that *X*

⁶ 'Only' here may seem a bit strong: but I take it that it is not an unreasonable demand on an account of his doing *the* right thing in *S* that his action is uniquely singled out from the alternatives. A further detail: one should strictly say, not that (1) implies a statement of the form (2), but that (1) implies *that there is* a true statement of that form.

would do the right thing only if he kept his promise; that keeping his promise involve (or consist in) doing A; that several other people would, as a matter of fact, keep their promises (and for the right reasons) if and only if X did not do A. There are all sorts of situations in which this sort of thing would be true: thus it might be the case that an effect of X's doing A would be to provide some inducement to these others which would lead them to break promises which otherwise they would have kept. Thus a non-consequentialist can hold both that it is a better state of affairs in which more people keep their promises, and that the right thing for X to do is something which brings it about that fewer promises are kept. Moreover, it is very obvious what view of things goes with holding that. It is one in which, even though from some abstract point of view one state of affairs is better than another, it does not follow that a given agent should regard it as his business to bring it about, even though it is open to him to do so. More than that, it might be that he could not properly regard it as his business. If the goodness of the world were to consist in people's fulfilling their obligations, it would by no means follow that a given agent should regard it as his business to bring it about, even though it is open to him to do so. More than that, it might be that he could not properly regard it as his business. If the goodness of the world were to consist in people's fulfilling their obligations, it would by no means follow that one of my obligations was to bring it about that other people kept their obligations.

Of course, no sane person could really believe that the goodness of the world just consisted in people keeping their obligations. But that is just an example, to illustrate the point that under non-consequentialism (3) does not, as one might expect, follow from (1). Thus even allowing some actions to have intrinsic value, we can still distinguish consequentialism. A consequentialist view, then, is one in which a statement of the form (2) follows from a statement of the form (1). A non-consequentialist view is one in which this is not so—not even when the (2)-statement takes the special form of (3).

This is not at all to say that the alternative to consequentialism is that one has to accept that there are some actions which one should always do, or again some which one should never do,

whatever the consequences: this is a much stronger position than any involved, as I have defined the issues, in the denial of consequentialism. All that is involved, on the present account, in the denial of consequentialism, is that with respect to some type of action, there are some situations in which that would be the right thing to do, even though the state of affairs produced by one's doing that would be worse than some other state of affairs accessible to one. The claim that there is a type of action which is right *whatever the consequences* can be put by saying that with respect to some type of action, assumed as being adequately specified, then *whatever* the situation may (otherwise) be, that will be the right thing to do, *whatever* other state of affairs might be accessible to one, however much better it might be than the state of affairs produced by one's doing this action.

If that somewhat Moorean formulation has not hopelessly concealed the point, it will be seen that this second position—the *whatever the consequences* position—is very much stronger than the first, the mere rejection of consequentialism. It is perfectly consistent, and it might be thought a mark of sense, to believe, while not being a consequentialist, that there was no type of action which satisfied this second condition: that if an adequate (and non-question-begging) specification of a type of action has been given in advance, it is always possible to think of some situation in which the consequences of doing the action so specified would be so awful that it would be right to do something else.

Of course, one might think that there just *were* some types of action which satisfied this condition; though it seems to me obscure how one could have much faith in a list of such actions unless one supposed that it had supernatural warrant. Alternatively, one might think that while logically there was a difference between the two positions, in social and psychological fact they came to much the same thing, since so soon (it might be claimed) as people give up thinking in terms of certain things being right or wrong *whatever the consequences*, they turn to thinking in purely consequential terms. This might be offered as a very general proposition about human thought, or (more plausibly) as a sociological proposition about certain situations of social change, in which utilitarianism (in particular) looks the only coherent

alternative to a dilapidated set of values. At the level of language, it is worth noting that the use of the word '*absolute*' mirrors, and perhaps also assists, this association: the claim that no type of action is 'absolutely right'—leaving aside the sense in which it means that the rightness of anything depends on the value-system of a society (the confused doctrine of relativism)—can mean either that no type of action is right-whatever-its-consequences, or, alternatively, that 'it all depends on the consequences', that is, in each case the decision whether an action is right is determined by its consequences.

A particular sort of psychological connection—or, in an old-fashioned use of the term, a 'moral' connection—between the two positions might be found in this. If people do not regard certain things as 'absolutely out', then they are prepared to start thinking about extreme situations in which what would otherwise be out might, exceptionally, be justified. They will, if they are to get clear about what they believe, be prepared to compare different extreme situations and ask what action would be justified in them. But once they have got used to that, their inhibitions about thinking of everything in consequential terms disappear: the difference between the extreme situations and the less extreme presents itself no longer as a difference between the exceptional and the usual, but between the greater and the less—and the consequential thoughts one was prepared to deploy in the greater it may seem quite irrational not to deploy in the less. *A fortiori*, someone might say: but he would have already had to complete this process to see it as a case of *a fortiori*.

One could regard this process of adaptation to consequentialism, moreover, not merely as a blank piece of psychological association, but as concealing a more elaborate structure of thought. One might have the idea that the *unthinkable* was itself a moral category; and in more than one way. It could be a feature of a man's moral outlook that he regarded certain courses of action as unthinkable, in the sense that he would not entertain the idea of doing them: and the witness to that might, in many cases, be that they simply would not come into his head. Entertaining certain alternatives, regarding them indeed as *alternatives*, is itself something that he regards as dishonourable or morally absurd.

But, further, he might equally find it unacceptable to consider what to do in certain conceivable situations. Logically, or indeed empirically conceivable they may be, but they are not to him morally conceivable, meaning by that that their occurrence as situations presenting him with a choice would represent not a special problem in his moral world, but something that lay beyond its limits. For him, there are certain situations so monstrous that the idea that the processes of moral rationality could yield an answer in them is insane: they are situations which so transcend in enormity the human business of moral deliberation that from a moral point of view it cannot matter any more what happens. Equally, for him, to spend time thinking what one would decide if one were in such a situation is also insane, if not merely frivolous.

For such a man, and indeed for anyone who is prepared to take him seriously, the demand, in Herman Kahn's words, to *think the unthinkable* is not an unquestionable demand of rationality, set against a cowardly or inert refusal to follow out one's moral thoughts. Rationality he sees as a demand not merely on him, but on the situations in, and about, which he has to think; unless the environment reveals minimum sanity, it is insanity to carry the decorum of sanity into it. Consequentialist rationality, however, and in particular utilitarian rationality, has no such limitations: making the best of a bad job is one of its maxims, and it will have something to say even on the difference between massacring seven million, and massacring seven million and one.

There are other important questions about the idea of the morally unthinkable which we cannot pursue here. Here we have been concerned with the role it might play in someone's connecting, by more than a mistake, the idea that there was nothing which was right whatever the consequences, and the different idea that everything depends on consequences. While someone might, in this way or another, move from one of those ideas to the other, it is very important that the two ideas are different: especially important in a world where we have lost traditional reasons for resisting the first idea, but have more than enough reasons for fearing the second.

2. NEGATIVE RESPONSIBILITY: AND TWO EXAMPLES

Although I have defined a state of affairs being *accessible* to an agent in terms of the actions which are *available* to him,⁷ nevertheless it is the former notion which is really more important for consequentialism. Consequentialism is basically indifferent to whether a state of affairs consists in what I do, or is produced by what I do, where that notion is itself wide enough to include, for instance, situations in which other people do things which I have made them do, or allowed them to do, or encouraged them to do, or given them a chance to do. All that consequentialism is interested in is the idea of these doings being *consequences* of what I do, and that is a relation broad enough to include the relations just mentioned, and many others.

Just what the relation is, is a different question, and at least as obscure as the nature of its relative, cause and effect. It is not a question I shall try to pursue; I will rely on cases where I suppose that any consequentialist would be bound to regard the situations in question as consequences of what the agent does. There are cases where the supposed consequences stand in a rather remote relation to the action, which are sometimes difficult to assess from a practical point of view, but which raise no very interesting question for the present enquiry. The more interesting points about consequentialism lie rather elsewhere. There are certain situations in which the causation of the situation, the relation it has to what I do, is in no way remote or problematic in itself, and entirely justifies the claim that the situation is a consequence of what I do: for instance, it is quite clear, or reasonably clear, that if I do a certain thing, this situation will come about, and if I do not, it will not. So from a consequentialist point of view it goes into the calculation of consequences along with any other state of affairs accessible to me. Yet from some, at least, non-consequentialist points of view, there is a vital difference between some such situations and others: namely, that in some a vital link in the production of the eventual outcome is provided by *someone else's* doing something. But for consequentialism, all casual connections

⁷ See last section, pp. 24-5.

are on the same level, and it makes no difference, so far as that goes, whether the causation of a given state of affairs lies through another agent, or not.

Correspondingly, there is no relevant difference which consists *just* in one state of affairs being brought about by me, without intervention of other agents, and another being brought about through the intervention of other agents; although some genuinely causal differences involving a difference of value may correspond to that (as when, for instance, the other agents derive pleasure or pain from the transaction), that kind of difference will already be included in the specification of the state of affairs to be produced. Granted that the states of affairs have been adequately described in causally and evaluatively relevant terms, it makes no further comprehensible difference who produces them. It is because consequentialism attaches value ultimately to states of affairs, and its concern is with what states of affairs the world contains, that it essentially involves the notion of *negative responsibility*: that if I am ever responsible for anything, then I must be just as much responsible for things that I allow or fail to prevent, as I am for things that I myself, in the more everyday restricted sense, bring about.⁸ Those things also must enter my deliberations, as a responsible moral agent, on the same footing. What matters is what states of affairs the world contains, and so what matters with respect to a given action is what comes about if it is done, and what comes about if it is not done, and those are questions not intrinsically affected by the nature of the causal linkage, in particular by whether the outcome is partly produced by other agents.

The strong doctrine of negative responsibility flows directly from consequentialism's assignment of ultimate value to states of affairs. Looked at from another point of view, it can be seen also as a special application of something that is favoured in many moral outlooks not themselves consequentialist—something which,

⁸ This is a fairly modest sense of 'responsibility', introduced merely by one's ability to reflect on, and decide, what one ought to do. This presumably escapes Smart's ban (p. 54) on the notion of 'the responsibility' as 'a piece of metaphysical nonsense'—his remarks seem to be concerned solely with situations of interpersonal blame.

indeed, some thinkers have been disposed to regard as the essence of morality itself: a principle of impartiality. Such a principle will claim that there can be no relevant difference from a moral point of view which consists just in the fact, not further explicable in general terms, that benefits or harms accrue to one person rather than to another—'it's me' can never in itself be a morally comprehensible reason.⁹ This principle, familiar with regard to the reception of harms and benefits, we can see consequentialism as extending to their production: from the moral point of view, there is no comprehensible difference which consists just in my bringing about a certain outcome rather than someone else's producing it. That the doctrine of negative responsibility represents in this way the extreme of impartiality, and abstracts from the identity of the agent, leaving just a locus of causal intervention in the world—that fact is not merely a surface paradox. It helps to explain why consequentialism can seem to some to express a more serious attitude than non-consequentialist views, why part of its appeal is to a certain kind of high-mindedness. Indeed, that is part of what is wrong with it.

For a lot of the time so far we have been operating at an exceedingly abstract level. This has been necessary in order to get clearer in general terms about the differences between consequentialist and other outlooks, an aim which is important if we want to know what features of them lead to what results for our thought. Now, however, let us look more concretely at two examples, to see what utilitarianism might say about them, what we might say about utilitarianism and, most importantly of all, what would be implied by certain ways of thinking about the situations. The examples are inevitably schematized, and they are open to the objection that they beg as many questions as they illuminate. There are two ways in particular in which examples in moral philosophy tend to beg important questions. One is that, as presented, they arbitrarily cut off and restrict the range of alternative courses of action—this objection might particularly be made against the first of my two examples. The second is that they

⁹ There is a tendency in some writers to suggest that it is not a comprehensible reason at all. But this, I suspect, is due to the overwhelming importance those writers ascribe to the moral point of view.

inevitably present one with the situation as a going concern, and cut off questions about how the agent got into it, and correspondingly about moral considerations which might flow from that: this objection might perhaps specially arise with regard to the second of my two situations. These difficulties, however, just have to be accepted, and if anyone finds these examples crippling defective in this sort of respect, then he must in his own thought rework them in richer and less question-begging form. If he feels that no presentation of any imagined situation can ever be other than misleading in morality, and that there can never be any substitute for the concrete experienced complexity of actual moral situations, then this discussion, with him, must certainly grind to a halt: but then one may legitimately wonder whether every discussion with him about conduct will not grind to a halt, including any discussion about the actual situations, since discussion about how one would think and feel about situations somewhat different from the actual (that is to say, situations to that extent imaginary) plays an important role in discussion of the actual.

(1) George, who has just taken his Ph.D. in chemistry, finds it extremely difficult to get a job. He is not very robust in health, which cuts down the number of jobs he might be able to do satisfactorily. His wife has to go out to work to keep them, which itself causes a great deal of strain, since they have small children and there are severe problems about looking after them. The results of all this, especially on the children, are damaging. An older chemist, who knows about this situation, says that he can get George a decently paid job in a certain laboratory, which pursues research into chemical and biological warfare. George says that he cannot accept this, since he is opposed to chemical and biological warfare. The older man replies that he is not too keen on it himself, come to that, but after all George's refusal is not going to make the job or the laboratory go away; what is more, he happens to know that if George refuses the job, it will certainly go to a contemporary of George's who is not inhibited by any such scruples and is likely if appointed to push along the research with greater zeal than George would. Indeed, it is not merely concern for George and his family, but (to speak frankly and in confidence) some alarm about this other man's excess of zeal, which has led the

older man to offer to use his influence to get George the job . . . George's wife, to whom he is deeply attached, has views (the details of which need not concern us) from which it follows that at least there is nothing particularly wrong with research into CBW. What should he do?

(2) Jim finds himself in the central square of a small South American town. Tied up against the wall are a row of twenty Indians, most terrified, a few defiant, in front of them several armed men in uniform. A heavy man in a sweat-stained khaki shirt turns out to be the captain in charge and, after a good deal of questioning of Jim which establishes that he got there by accident while on a botanical expedition, explains that the Indians are a random group of the inhabitants who, after recent acts of protest against the government, are just about to be killed to remind other possible protestors of the advantages of not protesting. However, since Jim is an honoured visitor from another land, the captain is happy to offer him a guest's privilege of killing one of the Indians himself. If Jim accepts, then as a special mark of the occasion, the other Indians will be let off. Of course, if Jim refuses, then there is no special occasion, and Pedro here will do what he was about to do when Jim arrived, and kill them all. Jim, with some desperate recollection of schoolboy fiction, wonders whether, if he got hold of a gun, he could hold the captain, Pedro, and the rest of the soldiers to threat, but it is quite clear from the set-up that nothing of that kind is going to work: any attempt at that sort of thing will mean that all the Indians will be killed, and himself. The men against the wall, and the other villagers, understand the situation, and are obviously begging him to accept. What should he do?

To these dilemmas, it seems to me that utilitarianism replies, in the first case, that George should accept the job, and in the second, that Jim should kill the Indian. Not only does utilitarianism give these answers but, if the situations are essentially as described and there are no further special factors, it regards them, it seems to me, as *obviously* the right answers. But many of us would certainly wonder whether, in (1), that could possibly be the right answer at all; and in the case of (2), even one who came to think that perhaps that was the answer, might well wonder whether it was obviously the answer. Nor is it just a question of the rightness

or obviousness of these answers. It is also a question of what sort of considerations come into finding the answer. A feature of utilitarianism is that it cuts out a kind of consideration which for some others makes a difference to what they feel about such cases: a consideration involving the idea, as we might first and very simply put it, that each of us is specially responsible for what *he* does, rather than for what other people do. This is an idea closely connected with the value of integrity. It is often suspected that utilitarianism, at least in its direct forms, makes integrity as a value more or less unintelligible. I shall try to show that this suspicion is correct. Of course, even if that is correct, it would not necessarily follow that we should reject utilitarianism; perhaps, as utilitarians sometimes suggest, we should just forget about integrity, in favour of such things as a concern for the general good. However, if I am right, we cannot merely do that, since the reason why utilitarianism cannot understand integrity is that it cannot coherently describe the relations between a man's projects and his actions.

3. TWO KINDS OF REMOTER EFFECT

A lot of what we have to say about this question will be about the relations between my projects and other people's projects. But before we get on to that, we should first ask whether we are assuming too hastily what the utilitarian answers to the dilemmas will be. In terms of more direct effects of the possible decisions, there does not indeed seem much doubt about the answer in either case; but it might be said that in terms of more remote or less evident effects counterweights might be found to enter the utilitarian scales. Thus the effect on George of a decision to take the job might be invoked, or its effect on others who might know of his decision. The possibility of there being more beneficent labours in the future from which he might be barred or disqualified, might be mentioned; and so forth. Such effects—in particular, possible effects on the agent's character, and effects on the public at large—are often invoked by utilitarian writers dealing with problems about lying or promise-breaking, and some similar considerations might be invoked here.

There is one very general remark that is worth making about

arguments of this sort. The certainty that attaches to these hypotheses about possible effects is usually pretty low; in some cases, indeed, the hypothesis invoked is so implausible that it would scarcely pass if it were not being used to deliver the respectable moral answer, as in the standard fantasy that one of the effects of one's telling a particular lie is to weaken the disposition of the world at large to tell the truth. The demands on the certainty or probability of these beliefs as beliefs about particular actions are much milder than they would be on beliefs favouring the unconventional course. It may be said that this is as it should be, since the presumption must be in favour of the conventional course: but that scarcely seems a *utilitarian* answer, unless utilitarianism has already taken off in the direction of not applying the consequences to the particular act at all.

Leaving aside that very general point, I want to consider now two types of effect that are often invoked by utilitarians, and which might be invoked in connection with these imaginary cases. The attitude or tone involved in invoking these effects may sometimes seem peculiar; but that sort of peculiarity soon becomes familiar in utilitarian discussions, and indeed it can be something of an achievement to retain a sense of it.

First, there is the psychological effect on the agent. Our descriptions of these situations have not so far taken account of how George or Jim will be after they have taken the one course or the other; and it might be said that if they take the course which seemed at first the utilitarian one, the effects on them will be in fact bad enough and extensive enough to cancel out the initial utilitarian advantages of that course. Now there is one version of this effect in which, for a utilitarian, some confusion must be involved, namely that in which the agent feels bad, his subsequent conduct and relations are crippled, and so on, *because he thinks that he has done the wrong thing*—for if the balance of outcomes was as it appeared to be *before* invoking this effect, then he has not (from the utilitarian point of view) done the wrong thing. So that version of the effect, for a rational and utilitarian agent, could not possibly make any difference to the assessment of right and wrong. However, perhaps he is not a thoroughly rational agent, and is disposed to have bad feelings, whichever he decided to do. Now

such feelings, which are from a strictly utilitarian point of view irrational—nothing, a utilitarian can point out, is advanced by having them—cannot, consistently, have any great weight in a utilitarian calculation. I shall consider in a moment an argument to suggest that they should have no weight at all in it. But short of that, the utilitarian could reasonably say that such feelings should not be encouraged, even if we accept their existence, and that to give them a lot of weight is to encourage them. Or, at the very best, even if they are straightforwardly and without any discount to be put into the calculation, their weight must be small: they are after all (and at best) one man's feelings.

That consideration might seem to have particular force in Jim's case. In George's case, his feelings represent a larger proportion of what is to be weighed, and are more commensurate in character with other items in the calculation. In Jim's case, however, his feelings might seem to be of very little weight compared with other things that are at stake. There is a powerful and recognizable appeal that can be made on this point: as that a refusal by Jim to do what he has been invited to do would be a kind of self-indulgent squeamishness. That is an appeal which can be made by other than utilitarians—indeed, there are some uses of it which cannot be consistently made by utilitarians, as when it essentially involves the idea that there is something dishonourable about such self-indulgence. But in some versions it is a familiar, and it must be said a powerful, weapon of utilitarianism. One must be clear, though, about what it can and cannot accomplish. The most it can do, so far as I can see, is to invite one to consider how seriously, and for what reasons, one feels that what one is invited to do is (in these circumstances) wrong, and, in particular, to consider that question from the utilitarian point of view. When the agent is not seeing the situation from a utilitarian point of view, the appeal cannot force him to do so; and if he does come round to seeing it from a utilitarian point of view, there is virtually nothing left for the appeal to do. If he does not see it from a utilitarian point of view, he will not see his resistance to the invitation, and the unpleasant feelings he associates with accepting it, *just* as disagreeable experiences of his; they figure rather as emotional expressions of a thought that to accept would be wrong. He may

be asked, as by the appeal, to consider whether he is right, and indeed whether he is fully serious, in thinking that. But the assertion of the appeal, that he is being self-indulgently squeamish, will not itself answer that question, or even help to answer it, since it essentially tells him to regard his feelings just as unpleasant experiences of his, and he cannot, by doing that, answer the question they pose when they are precisely not so regarded, but are regarded as indications¹⁰ of what he thinks is right and wrong. If he does come round fully to the utilitarian point of view then of course he will regard these feelings just as unpleasant experiences of his. And once Jim—at least—has come to see them in that light, there is nothing left for the appeal to do, since *of course* his feelings, so regarded, are of virtually no weight at all in relation to the other things at stake. The 'squeamishness' appeal is not an argument which adds in a hitherto neglected consideration. Rather, it is an invitation to consider the situation, and one's own feelings, from a utilitarian point of view.

The reason why the squeamishness appeal can be very unsettling, and one can be unnerved by the suggestion of self-indulgence in going against utilitarian considerations, is not that we are utilitarians who are uncertain what utilitarian value to attach to our moral feelings, but that we are partially at least not utilitarians, and cannot regard our moral feelings merely as objects of utilitarian value. Because our moral relation to the world is partly given by such feelings, and by a sense of what we can or cannot 'live with', to come to regard those feelings from a purely utilitarian point of view, that is to say, as happenings outside one's moral self, is to lose a sense of one's moral identity; to lose, in the most literal way, one's integrity. At this point utilitarianism alienates one from one's moral feelings; we shall see a little later how, more basically, it alienates one from one's actions as well.

If, then, one is really going to regard one's feelings from a strictly utilitarian point of view, Jim should give very little weight at all to his; it seems almost indecent, in fact, once one has taken

¹⁰ On the non-cognitivist meta-ethic in terms of which Smart presents his utilitarianism, the term 'indications' here would represent an understatement.

that point of view, to suppose that he should give any at all. In George's case one might feel that things were slightly different. It is interesting, though, that one reason why one might think that—namely that one person principally affected is his wife—is very dubiously available to a utilitarian. George's wife has some reason to be interested in George's integrity and his sense of it; the Indians, quite properly, have no interest in Jim's. But it is not at all clear how utilitarianism would describe that difference.

There is an argument, and a strong one, that a strict utilitarian should give not merely small extra weight, in calculations of right and wrong, to feelings of this kind, but that he should give absolutely no weight to them at all. This is based on the point, which we have already seen, that if a course of action is, before taking these sorts of feelings into account, utilitarianly preferable, then bad feelings about that kind of action will be from a utilitarian point of view irrational. Now it might be thought that even if that is so, it would not mean that in a utilitarian calculation such feelings should not be taken into account; it is after all a well-known boast of utilitarianism that it is a realistic outlook which seeks the best in the world as it is, and takes any form of happiness or unhappiness into account. While a utilitarian will no doubt seek to diminish the incidence of feelings which are utilitarianly irrational—or at least of disagreeable feelings which are so—he might be expected to take them into account while they exist. This is without doubt classical utilitarian doctrine, but there is good reason to think that utilitarianism cannot stick to it without embracing results which are startlingly unacceptable and perhaps self-defeating.

Suppose that there is in a certain society a racial minority. Considering merely the ordinary interests of the other citizens, as opposed to their sentiments, this minority does no particular harm; we may suppose that it does not confer any very great benefits either. Its presence is in those terms neutral or mildly beneficial. However, the other citizens have such prejudices that they find the sight of this group, even the knowledge of its presence, very disagreeable. Proposals are made for removing in some way this minority. If we assume various quite plausible things (as that programmes to change the majority sentiment are

likely to be protracted and ineffective) then even if the removal would be unpleasant for the minority, a utilitarian calculation might well end up favouring this step, especially if the minority were a rather small minority and the majority were very severely prejudiced, that is to say, were made very severely uncomfortable by the presence of the minority.

A utilitarian might find that conclusion embarrassing; and not merely because of its nature, but because of the grounds on which it is reached. While a utilitarian might be expected to take into account certain other sorts of consequences of the prejudice, as that a majority prejudice is likely to be displayed in conduct disagreeable to the minority, and so forth, he might be made to wonder whether the unpleasant experiences of the prejudiced people should be allowed, *merely as such*, to count. If he does count them, merely as such, then he has once more separated himself from a body of ordinary moral thought which he might have hoped to accommodate; he may also have started on the path of defeating his own view of things. For one feature of these sentiments is that they are from the utilitarian point of view itself irrational, and a thoroughly utilitarian person would either not have them, or if he found that he did tend to have them, would himself seek to discount them. Since the sentiments in question are such that a rational utilitarian would discount them in himself, it is reasonable to suppose that he should discount them in his calculations about society; it does seem quite unreasonable for him to give just as much weight to feelings—considered just in themselves, one must recall, as experiences of those that have them—which are essentially based on views which are from a utilitarian point of view irrational, as to those which accord with utilitarian principles. Granted this idea, it seems reasonable for him to rejoin a body of moral thought in other respects congenial to him, and discount those sentiments, just considered in themselves, totally, on the principle that no pains or discomforts are to count in the utilitarian sum which their subjects have just because they hold views which are by utilitarian standards irrational. But if he accepts that, then in the cases we are at present considering no extra weight at all can be put in for bad feelings of George or Jim about their choices, if those choices are,

leaving out those feelings, on the first round utilitarianly rational.

The psychological effect on the agent was the first of two general effects considered by utilitarians, which had to be discussed. The second is in general a more substantial item, but it need not take so long, since it is both clearer and has little application to the present cases. This is the *precedent effect*. As Burke rightly emphasized, this effect can be important: that one morally *can* do what someone has actually done, is a psychologically effective principle, if not a deontically valid one. For the effect to operate, obviously some conditions must hold on the publicity of the act and on such things as the status of the agent (such considerations weighed importantly with Sir Thomas More); what these may be will vary evidently with circumstances.

In order for the precedent effect to make a difference to a utilitarian calculation, it must be based upon a confusion. For suppose that there is an act which would be the best in the circumstances, except that doing it will encourage by precedent other people to do things which will not be the best things to do. Then the situation of those other people must be relevantly different from that of the original agent; if it were not, then in doing the same as what would be the best course for the original agent, they would necessarily do the best thing themselves. But if the situations are in this way relevantly different, it must be a confused perception which takes the first situation, and the agent's course in it, as an adequate precedent for the second.

However, the fact that the precedent effect, if it really makes a difference, is in this sense based on a confusion, does not mean that it is not perfectly real, nor that it is to be discounted: social effects are by their nature confused in this sort of way. What it does emphasize is that calculations of the precedent effect have got to be realistic, involving considerations of how people are actually likely to be influenced. In the present examples, however, it is very implausible to think that the precedent effect could be invoked to make any difference to the calculation. Jim's case is extraordinary enough, and it is hard to imagine who the recipients of the effect might be supposed to be; while George is not in a sufficiently public situation or role for the question to arise in that

form, and in any case one might suppose that the motivations of others on such an issue were quite likely to be fixed one way or another already.

No appeal, then, to these other effects is going to make a difference to what the utilitarian will decide about our examples. Let us now look more closely at the structure of those decisions.

4. INTEGRITY

The situations have in common that if the agent does not do a certain disagreeable thing, someone else will, and in Jim's situation at least the result, the state of affairs after the other man has acted, if he does, will be worse than after Jim has acted, if Jim does. The same, on a smaller scale, is true of George's case. I have already suggested that it is inherent in consequentialism that it offers a strong doctrine of negative responsibility: if I know that if I do X , O_1 will eventuate, and if I refrain from doing X , O_2 will, and that O_2 is worse than O_1 , then I am responsible for O_2 if I refrain voluntarily from doing X . 'You could have prevented it', as will be said, and truly, to Jim, if he refuses, by the relatives of the other Indians. (I shall leave the important question, which is to the side of the present issue, of the obligations, if any, that nest round the word 'know': how far does one, under utilitarianism, have to research into the possibilities of maximally beneficent action, including prevention?)

In the present cases, the situation of O_2 includes another agent bringing about results worse than O_1 . So far as O_2 has been identified up to this point—merely as the worse outcome which will eventuate if I refrain from doing X —we might equally have said that what that other brings about is O_2 ; but that would be to underdescribe the situation. For what occurs if Jim refrains from action is not solely twenty Indians dead, but *Pedro's killing twenty Indians*, and that is not a result which Pedro brings about, though the death of the Indians is. We can say: what one does is not included in the outcome of what one does, while what another does can be included in the outcome of what one does. For that to be so, as the terms are now being used, only a very weak condition

has to be satisfied: for Pedro's killing the Indians to be the outcome of Jim's refusal, it only has to be causally true that if Jim had not refused, Pedro would not have done it.

That may be enough for us to speak, in some sense, of Jim's responsibility for that outcome, if it occurs; but it is certainly not enough, it is worth noticing, for us to speak of Jim's *making* those things happen. For granted this way of their coming about, he could have made them happen only by making Pedro shoot, and there is no acceptable sense in which his refusal makes Pedro shoot. If the captain had said on Jim's refusal, 'you leave me with no alternative', he would have been lying, like most who use that phrase. While the deaths, and the killing, may be the outcome of Jim's refusal, it is misleading to think, in such a case, of Jim having an *effect* on the world through the medium (as it happens) of Pedro's acts; for this is to leave Pedro out of the picture in his essential role of one who has intentions and projects, projects for realizing which Jim's refusal would leave an opportunity. Instead of thinking in terms of supposed effects of Jim's projects on Pedro, it is more revealing to think in terms of the effects of Pedro's projects on Jim's decision. This is the direction from which I want to criticize the notion of negative responsibility.

There are of course other ways in which this notion can be criticized. Many have hoped to discredit it by insisting on the basic moral relevance of the distinction between action and inaction, between intervening and letting things take their course. The distinction is certainly of great moral significance, and indeed it is not easy to think of any moral outlook which could get along without making some use of it. But it is unclear, both in itself and in its moral applications, and the unclarity is of a kind which precisely cause it to give way when, in very difficult cases, weight has to be put on it. There is much to be said in this area, but I doubt whether the sort of dilemma we are considering is going to be resolved by a simple use of this distinction. Again, the issue of negative responsibility can be pressed on the question of how limits are to be placed on one's apparently boundless obligation, implied by utilitarianism, to improve the world. Some answers are needed to that, too—and answers which stop short of relapsing into the bad faith of supposing that one's responsibilities could be

adequately characterized just by appeal to one's role.¹¹ But, once again, while that is a real question, it cannot be brought to bear directly on the present kind of case, since it is hard to think of anyone supposing that in Jim's case it would be an adequate response for him to say that it was none of his business.

What projects does a utilitarian agent have? As a utilitarian, he has the general project of bringing about maximally desirable outcomes; how he is to do this at any given moment is a question of what causal levers, so to speak, are at that moment within reach. The desirable outcomes, however, do not just consist of agents carrying out *that* project; there must be other more basic or lower-order projects which he and other agents have, and the desirable outcomes are going to consist, in part, of the maximally harmonious realization of those projects ('in part', because one component of a utilitarianly desirable outcome may be the occurrence of agreeable experiences which are not the satisfaction of anybody's projects). Unless there were first-order projects, the general utilitarian project would have nothing to work on, and would be vacuous. What do the more basic or lower-order projects comprise? Many will be the obvious kinds of desires for things for oneself, one's family, one's friends, including basic necessities of life and, in more relaxed circumstances, objects of taste. Or there may be pursuits and interests of an intellectual, cultural, or creative character. I introduce those as a separate class not because the objects of them lie in a separate class and provide—as some utilitarians, in their churchy way, are fond of saying—'higher' pleasures. I introduce them separately because the agent's identification with them may be of a different order. It does not have to be: cultural and aesthetic interests just belong, for many, along with any other taste; but some people's commitment to these kinds of interests just is at once more thoroughgoing and serious than their pursuit of various objects of taste, while it is more individual and permeated with character than the desire for the necessities of life.

Beyond these, someone may have projects connected with his

¹¹ For some remarks bearing on this, see *Morality*, the section on 'Goodness and roles', and Cohen's article there cited. [Williams's book *Morality* was published by Harper & Row in 1972—Ed.]

support of some cause: Zionism, for instance, or the abolition of chemical and biological warfare. Or there may be projects which flow from some more general disposition towards human conduct and character, such as a hatred of injustice, or of cruelty, or of killing.

It may be said that this last sort of disposition and its associated project do not count as (logically) 'lower-order' relative to the higher-order project of maximizing desirable outcomes; rather, it may be said, it is itself a 'higher-order' project. The vital question is not, however, how it is to be classified, but whether it and similar projects are to count among the projects whose satisfaction is to be included in the maximizing sum and, correspondingly, as contributing to the agent's happiness. If the utilitarian says 'no' to that, then he is almost certainly committed to a version of utilitarianism as absurdly superficial and shallow as Benthamite versions have often been accused of being. For this project will be discounted, presumably, on the ground that it involves, in the specification of its object, the mention of other people's happiness or interests: thus it is the kind of project which (unlike the pursuit of food for myself) presupposes a reference to other people's projects. But that criterion would eliminate any desire at all which was not blankly and in the most straightforward sense egoistic.¹² Thus we should be reduced to frankly egoistic first-order projects and—for all essential purposes—the one second-order utilitarian project of maximally satisfying first-order projects. Utilitarianism has a tendency to slide in this direction, and to leave a vast hole in the range of human desires, between egoistic inclinations and necessities at one end, and impersonally benevolent happiness-management at the other. But the utilitarianism which has to leave this hole is the most primitive form, which offers a quite rudimentary account of desire. Modern versions of the theory are supposed to be neutral with regard to what sorts of things make people happy or what their projects are. Utilitarianism would do well then to acknowledge the evident fact that among the things that make people happy is not only making other people happy,

¹² On the subject of egoistic and non-egoistic desires, see 'Egoism and Altruism', in *Problems of the Self* (London: Cambridge University Press, 1973).

but being taken up or involved in any of a vast range of projects, or—if we waive the evangelical and moralizing associations of the word—commitments. One can be committed to such things as a person, a cause, an institution, a career, one's own genius, or the pursuit of danger.

Now none of these is itself the *pursuit of happiness*: by an exceedingly ancient platitude, it is not at all clear that there could be anything which was just that, or at least anything that had the slightest chance of being successful. Happiness, rather, requires being involved in, or at least content with, something else.¹³ It is not impossible for utilitarianism to accept that point: it does not have to be saddled with a naive and absurd philosophy of mind about the relation between desire and happiness. What it does have to say is that if such commitments are worth while, then pursuing the projects that flow from them, and realizing some of those projects, will make the person for whom they are worth while, happy. It may be that to claim that is still wrong: it may well be that a commitment can make sense to a man (can make sense of his life) without his supposing that it will make him *happy*.¹⁴ But that is not the present point; let us grant to utilitarianism that all worthwhile human projects must conduce, one way or another, to happiness. The point is that even if that is true, it does not follow, nor could it possibly be true, that those projects are themselves projects of pursuing happiness. One has to believe in, or at least want, or, quite minimally, be content with, other things, for there to be anywhere that happiness can come from.

Utilitarianism, then, should be willing to agree that its general aim of maximizing happiness does not imply that what everyone is doing is just pursuing happiness. On the contrary, people have to be pursuing other things. What those other things may be,

¹³ This does not imply that there is no such thing as the project of pursuing pleasure. Some writers who have correctly resisted the view that all desires are desires for pleasure, have given an account of pleasure so thoroughly adverbial as to leave it quite unclear how there could be a distinctively hedonist way of life at all. Some room has to be left for that, though there are important difficulties both in defining it and living it. Thus (particularly in the case of the very rich) it often has highly ritual aspects, apparently part of a strategy to counter boredom.

¹⁴ For some remarks on this possibility, see *Morality*, section on 'What is morality about?'

utilitarianism, sticking to its professed empirical stance, should be prepared just to find out. No doubt some possible projects it will want to discourage, on the grounds that their being pursued involves a negative balance of happiness to others: though even there, the unblinking accountant's eye of the strict utilitarian will have something to put in the positive column, the satisfactions of the destructive agent. Beyond that, there will be a vast variety of generally beneficent or at least harmless projects; and some, no doubt, will take the form not just of tastes or fancies, but of what I have called 'commitments'. It may even be that the utilitarian researcher will find that many of those with commitments, who have really identified themselves with objects outside themselves, who are thoroughly involved with other persons, or institutions, or activities or causes, are actually happier than those whose projects and wants are not like that. If so, that is an important piece of utilitarian empirical lore.

When I say 'happier' here, I have in mind the sort of consideration which any utilitarian would be committed to accepting: as for instance that such people are less likely to have a breakdown or commit suicide. Of course that is not all that is actually involved, but the point in this argument is to use to the maximum degree utilitarian notions, in order to locate a breaking-point in utilitarian thought. In appealing to this strictly utilitarian notion, I am being more consistent with utilitarianism than Smart is. In his struggles with the problem of the brain-electrode man, Smart (p. 22) commends the idea that 'happy' is a partly evaluative term, in the sense that we call 'happiness' those kinds of satisfaction which, as things are, we approve of. But *by what standard* is this surplus element of approval supposed, from a utilitarian point of view, to be allocated? There is no source for it, on a strictly utilitarian view, except further degrees of satisfaction, but there are none of those available, or the problem would not arise. Nor does it help to appeal to the fact that we dislike in prospect things which we like when we get there, for from a utilitarian point of view it would seem that the original dislike was merely irrational or based on an error. Smart's argument at this point seems to be embarrassed by a well-known utilitarian uneasiness, which comes from a feeling that it is not respectable to

ignore the 'deep', while not having anywhere left in human life to locate it.¹⁵

Let us now go back to the agent as utilitarian, and his higher-order project of maximizing desirable outcomes. At this level, he is committed only to that: what the outcome will actually consist of will depend entirely on the facts, on what persons with what projects and what potential satisfactions there are within calculable reach of the causal levers near which he finds himself. His own substantial projects and commitments come into it, but only as one lot among others—they potentially provide one set of satisfactions among those which he may be able to assist from where he happens to be. He is the agent of the satisfaction system who happens to be at a particular point at a particular time: in Jim's case, our man in South America. His own decisions as a utilitarian agent are a function of all the satisfactions which he can affect from where he is: and this means that the projects of others, to an indeterminately great extent, determine his decision.

This may be so either positively or negatively. It will be so positively if agents within the causal field of his decision have projects which are at any rate harmless, and so should be assisted. It will equally be so, but negatively, if there is an agent within the causal field whose projects are harmful, and have to be frustrated to maximize desirable outcomes. So it is with Jim and the soldier Pedro. On the utilitarian view, the undesirable projects of other people as much determine, in this negative way, one's decisions as the desirable ones do positively: if those people were not there, or had different projects, the causal nexus would be different, and it is the actual state of the causal nexus which determines the decision. The determination to an indefinite degree of my decisions by other people's projects is just another aspect of my unlimited responsibility to act for the best in a causal framework formed to a considerable extent by their projects.

The decision so determined is, for utilitarianism, the right decision. But what if it conflicts with some project of mine? This, the utilitarian will say, has already been dealt with: the satisfaction

¹⁵ One of many resemblances in spirit between utilitarianism and high-minded evangelical Christianity.

to you of fulfilling your project, and any satisfactions to others of your so doing, have already been through the calculating device and have been found inadequate. Now in the case of many sorts of projects, that is a perfectly reasonable sort of answer. But in the case of projects of the sort I have called 'commitments', those with which one is more deeply and extensively involved and identified, this cannot just by itself be an adequate answer, and there may be no adequate answer at all. For, to take the extreme sort of case, how can a man, a utilitarian agent, come to regard as one satisfaction among others, and a dispensable one, a project or attitude round which he has built his life, just because someone else's projects have so structured the causal scene that that is how the utilitarian sum comes out?

The point here is not, as utilitarians may hasten to say, that if the project or attitude is that central to his life, then to abandon it will be very disagreeable to him and great loss of utility will be involved. I have already argued in Section 3 that it is not like that; on the contrary, once he is prepared to look at it like that, the argument in any serious case is over anyway. The point is that he is identified with his actions as flowing from projects and attitudes which in some cases he takes seriously at the deepest level, as what his life is about (or, in some cases, this section of his life—seriousness is not necessarily the same as persistence). It is absurd to demand of such a man, when the sums come in from the utility network which the projects of others have in part determined, that he should just step aside from his own project and decision and acknowledge the decision which utilitarian calculation requires. It is to alienate him in a real sense from his actions and the source of his action in his own convictions. It is to make him into a channel between the input of everyone's projects, including his own, and an output of optimistic decision; but this is to neglect the extent to which *his* actions and *his* decisions have to be as the actions and decisions which flow from the projects and attitudes with which he is most closely identified. It is thus, in the most literal sense, an attack on his integrity.¹⁶

¹⁶ Interestingly related to these notions is the Socratic idea that courage is a virtue particularly connected with keeping a clear sense of what one regards as most important. They also centrally raise questions about the value of pride. Humility,

These sorts of considerations do not in themselves give solutions to practical dilemmas such as those provided by our examples; but I hope they help to provide other ways of thinking about them. In fact, it is not hard to see that in George's case, viewed from this perspective, the utilitarian solution would be wrong. Jim's case is different, and harder. But if (as I suppose) the utilitarian is probably right in this case, that is not to be found out just by asking the utilitarian's questions. Discussions of it—and I am not going to try to carry it further here—will have to take seriously the distinction between my killing someone, and its coming about because of what I do that someone else kills them: a distinction based, not so much on the distinction between action and inaction, as on the distinction between my projects and someone else's projects. At least it will have to start by taking that seriously, as utilitarianism does not; but then it will have to build out from there by asking why that distinction seems to have less, or a different, force in this case than it has in George's. One question here would be how far one's powerful objection to killing people just is, in fact, an application of a powerful objection to their being killed. Another dimension of that is the issue of how much it matters that the people at risk are actual, and there, as opposed to hypothetical, or future, or merely elsewhere.¹⁷

There are many other considerations that could come into such a question, but the immediate point of all this is to draw one particular contrast with utilitarianism: that to reach a grounded decision in such a case should not be regarded as a matter of just discounting one's reactions, impulses, and deeply held projects in the face of the pattern of utilities, nor yet merely adding them in—but in the first instance of trying to understand them.

as something beyond the real demand of correct self-appraisal, was specially a Christian virtue because it involved subservience to God. In a secular context it can only represent subservience to other men and their projects.

¹⁷ For a more general discussion of this issue see C. Fried, *An Anatomy of Values* (Cambridge, Mass.: Harvard University Press, 1970), Part Three.